Genomics 100 (2012) 72–80

Contents lists available at SciVerse ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Potential G-quadruplex formation at breakpoint regions of chromosomal translocations in cancer may explain their fragility

Vijeth K. Katapadi, Mridula Nambiar, Sathees C. Raghavan *

Department of Biochemistry, Indian Institute of Science, Bangalore-560 012, India

ARTICLE INFO

Article history:
Received 10 April 2012
Accepted 21 May 2012
Available online 30 May 2012

Keywords:
Genomic instability
Double-strand break
Altered DNA structures
DNA damage
Nonhomologous DNA end joining
Chromosomal translocation

ABSTRACT

Genetic alterations like point mutations, insertions, deletions, inversions and translocations are frequently found in cancers. Chromosomal translocations are one of the most common genomic aberrations associated with nearly all types of cancers especially leukemia and lymphoma. Recent studies have shown the role of non-B DNA structures in generation of translocations. In the present study, using various bioinformatic tools, we show the propensity of formation of different types of altered DNA structures near translocation breakpoint regions. In particular, we find close association between occurrence of G-quadruplex forming motifs and fragile regions in almost 70% of genes involved in rearrangements in lymphoid cancers. However, such an analysis did not provide any evidence for the occurrence of G-quadruplexes at the close vicinity of translocation breakpoint regions in nonlymphoid cancers. Overall, this study will help in the identification of novel non-B DNA targets that may be responsible for generation of chromosomal translocations in cancer.

1. Introduction

Cancers are typically associated with various kinds of genetic mutations which help in its progression and development. These alterations include point mutations, insertions, deletions, inversions and chromosomal translocations. A chromosomal translocation involves exchange of segments between two heterologous chromosomes [1]. Typically it results in the deregulation of important cellular genes as a result of its juxtaposition to enhancer/promoter elements, or in novel fusions [2–5]. It has been well studied that cancers of hematopoietic origin are characteristically associated with the presence of unique chromosomal translocations. Other neoplasia like sarcomas, myelomas and very recently the carcinomas, have also been shown to possess translocations and gene fusions, which are responsible for their pathophysiology [6–8]. The *BCL2-IgH* translocation is a genetic hallmark of follicular lymphoma and results in the up regulation of the anti-apoptotic, *BCL2* gene [9–11]. The Philadelphia chromosome is another common gene fusion between the *BCR* and *ABL* genes, associated with chronic myeloid leukemia, and leads to the formation of an activated tyrosine kinase, which disrupts the normal signaling cascade within a cell [1,12,13]. More than 500 chromosomal translocations have been identified so far, and many are yet to be uncovered, especially in carcinomas [2].

Conventionally, DNA in our genome exists in the B-form, exhibiting Watson–Crick hydrogen bonding. However, various non-B DNA structures like G-quadruplexes, triplexes, cruciforms, hairpins, Z-DNA etc. are slowly being recognized to play crucial roles in the regulation of many cellular processes (Fig. 1) [14–17]. The role of G-quadruplexes in transcription and telomeric stability is well characterized [18,19]. Very recently, G-quadruplex structures were shown to be responsible for the fragility of the *BCL2* mbr peak I region involved in the t(14;18) translocation in follicular lymphoma [20–22]. Presence of G-loop structures in the *c-MYC (MYC)* promoter region has also been suggested to impart fragility leading to translocations [23]. In addition, palindromic AT-rich repeats (PATRRs) are also shown to be involved in t(11;22)(q23;q11) and many other translocations [24,25]. PATRR11 on 11q23 can form a cruciform DNA structure in vitro as well [25].

Although chromosomal translocations were discovered more than half a century back, very little is known about the mechanism of its generation. Studies have identified genetic hotspots or fragile regions throughout the genome, which are more susceptible to breaks and are frequently involved in translocations [21]. Bioinformatics studies have suggested the propensity of formation of many non-B DNA structures throughout the genome [26]. However, their possible role in imparting fragility leading to chromosomal translocations has not been well studied.

In the present study, using a web based database approach and analytical tools we show a direct correlation between the occurrence of altered DNA structures and the cause of fragility at many of the translocation breakpoint regions. Further, we specifically determine the incidence of G-quadruplex DNA structures particularly in several fragile regions of lymphoid origin and identify several G-quadruplex DNA forming motifs in these regions. Hence, this study has attempted

* Corresponding author. Fax: +91 80 2360 0814.
E-mail address: sathees@biochem.iisc.ernet.in (S.C. Raghavan).

0888-7543/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygeno.2012.05.008

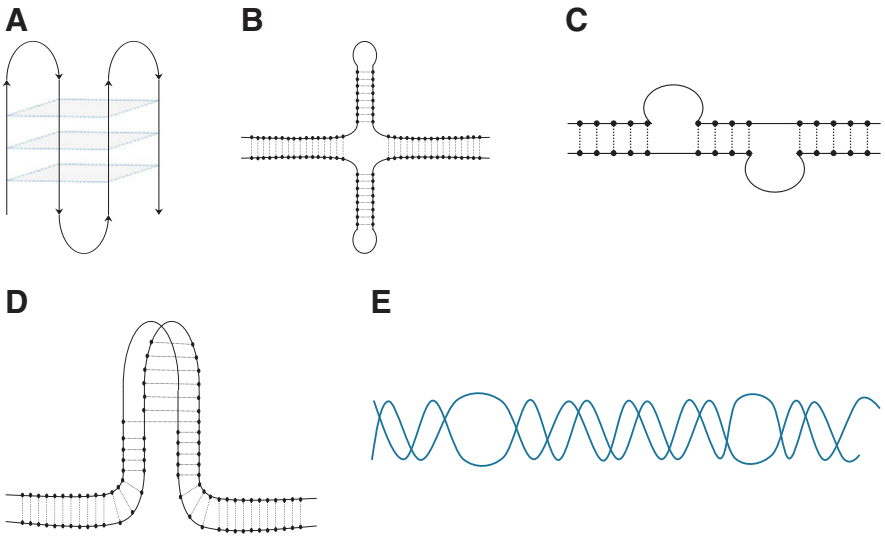


Fig. 1. Schematic representation of the different non-B DNA structures in the genome. A. G-quadruplex DNA (tetraplex) structures can be formed when at least four stretches of guanines are present in the DNA. Guanine tetrad formation is dependent on Hoogsteen hydrogen bonding at N7 position. A quadruplex can be either intramolecular (as shown here) or intermolecular based on the number of strands involved. B. Cruciform DNA can be formed by inverted repeats. The center of symmetry in each strand of the inverted repeat folds and adopts an intramolecular B-helix, which is usually capped by a single-stranded loop at the top. C. Slipped DNA structures are formed due to misalignment of direct repeats. D. Intramolecular triplex DNA (H-DNA) are three-stranded DNA structures and can be formed by polypurine–polypyrimidine sequences with mirror repeat symmetry. Hoogsteen hydrogen bonding is involved in the formation of triple helical DNA. E. Z DNA is a left-handed double-helical DNA structure with a zigzag pattern of base stacking. It is formed by the regions of alternating pyrimidine–purine, (YR–YR)n sequences at specific salt and moisture conditions.

to provide many potential non-B DNA targets, which will prove useful in understanding the mechanism of generation of several chromosomal translocations, which have remained unknown so far.

2. Results

2.1. Distribution of non-B DNA structures across species

It is believed that at least 2–5% of our genome may possess non-B form of DNA. Several altered DNA structures like G-quadruplexes, triplexes, cruciforms, Z DNA etc. have been speculated to be present in our genome (Fig. 1) [14–17]. Comparison of the different types of altered DNA structures found in the genomes across different mammalian species showed a similar pattern of distribution except in mouse genome, suggesting that the propensity to form altered DNA structures is comparable among the species (Suppl. Fig. 1) [26].

2.2. Non-B DNA structures and chromosomal translocations

The role of non-B DNA structures in genomic fragility during chromosomal translocations is an area of active investigation. Few evidences have been reported recently, which support the possible contribution of these structures towards generating genomic instability in cancers [21,23,24,27]. In order to test in silico, the occurrence of altered DNA structures in the genes known to be extensively involved in translocations, we mapped patient breakpoint regions from the data available in literature, using TICdb and COSMIC databases (data not shown). In addition, we also plotted various non-B DNA structures present at or near the patient breakpoint regions for genes involved in both lymphoid and non-lymphoid chromosomal translocations using non-B DB database (Figs. 2–4). We observed the presence of several non-canonical DNA structures within or near the major breakpoint region clusters in 10 of the 11 translocation associated genes studied like *RET*, *NCOA4*, *BCR*, *FGFR1*, *ABL1*, *E2A (TCF3)*, *CCDC6* etc. in lymphoid cancers (Fig. 3). Interestingly, we did not find the presence of any non-B DNA structure in the *PDGFRA*-exon 12 breakpoint region, suggesting an alternate mechanism for generation of the translocation (Fig. 3A). In the absence of any non-B DNA structure in this region, we can consider *PDGFRA*-exon 12 as a control for a region undergoing translocation by using a mechanism other than that proposed in the present study. Out of all the different types of altered structures, G-quadruplexes were found to be most frequently present near the mapped breakpoints in at least 13 breakpoint regions analyzed. In a fraction of the other cases like *BCR* intron 14, *CCDC6* intron 8, *ABL1* intron 1 and *NCOA4* intron 7, breakpoints were found near other altered structures like inverted and A phased repeats.

In case of genes involved in the non-lymphoid translocations, many of the patient breakpoint junctions are not well characterized. However, the major fragile sites are known and these were mapped along with the identified non-B DNA structure motifs for various genes. Results showed the propensity of several non-B DNA structures to be formed within or near the fragile sites in all the 7 breakpoint regions studied, which probably act as a trigger for undergoing breaks and further generation of translocations (Figs. 2, 4). Interestingly, in these cases, the fragile sites were flanked by different types of DNA structures and hence, the breaks could be contributed by formation of one or more non B DNA structures. Except for *ETV* and *TMPRSS2* genes, all other breakpoint regions studied have a more or less random distribution of different structures like Z-DNA, inverted and A-phased repeats around the patient breakpoints (Fig. 4).

2.3. G-quadruplexes during chromosomal translocations

Recent studies have shown the contribution of G-quadruplex structures in imparting genomic fragility during some of the most common translocations in follicular and Burkitt's lymphoma [20,21,23]. In the present study, we tried to validate the formation of such G-quadruplex structures at the *BCL2* mbr and c-*MYC (MYC)* breakpoint region by using previously described bioinformatic tools and found that we could indeed identify the G-quadruplex forming motifs in these genes (Fig. 3A, B). We further specifically looked into the possible connection between G-quadruplexes and genomic fragility in some of the fragile regions like *BCL1(CCND1)* MTC, *E2A(TCF3)*, *BCR* intron 14, *NCOA4* intron 7, *HOX11*, *ERG*, *FLI1*, *TMPRSS2* etc., which are associated with either lymphoid or non-lymphoid translocations. The G-quadruplex forming motifs were mapped both downstream and upstream of the breakpoint cluster region and the distance between them was calculated. Results showed the occurrence of G-quadruplex forming motifs near the patient breakpoint

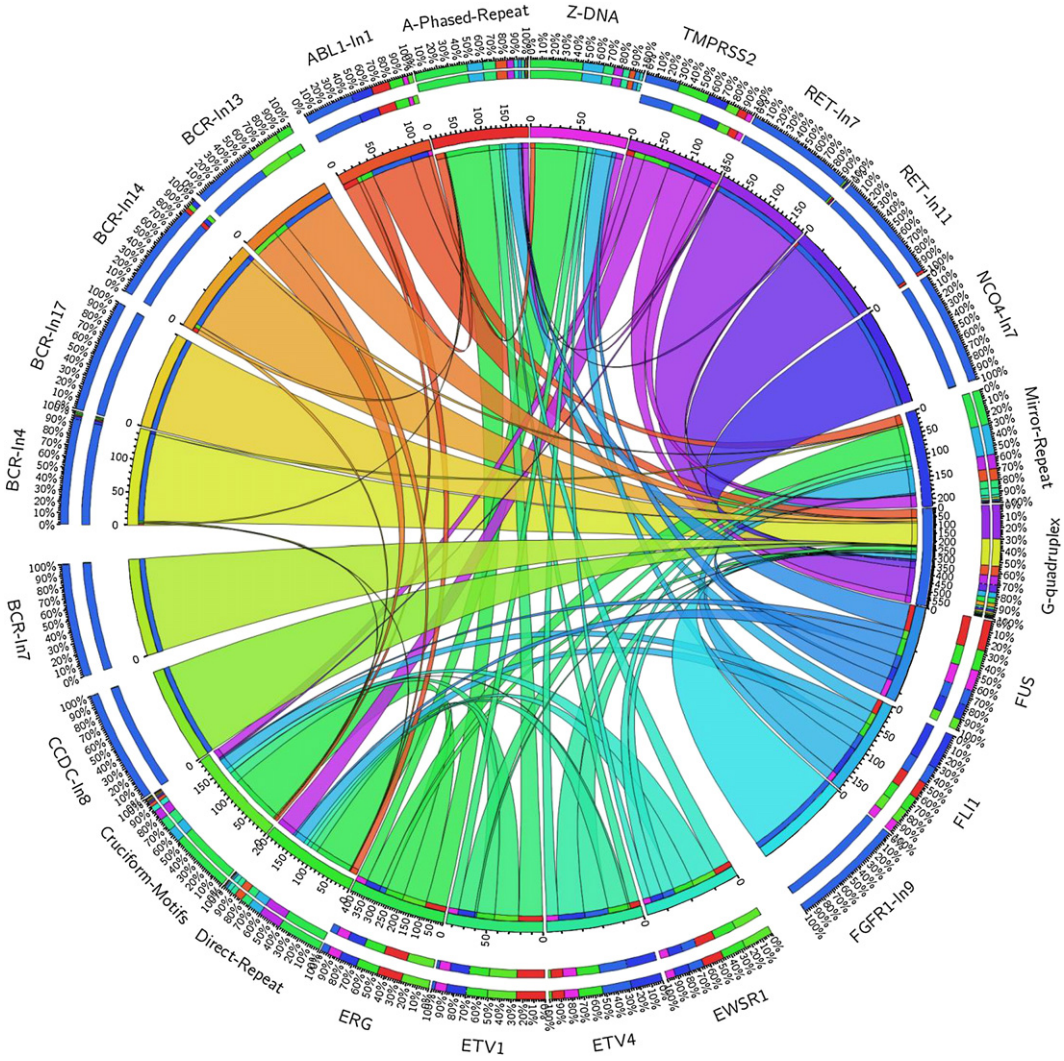


Fig. 2. Graphical representation of human genes involved in translocations of both lymphoid and non-lymphoid cancers, harboring different types of non-B DNA structures. The circos plot depicts different genes involved in human cancers in the inner ring and various non-B DNA structures in the outer ring. Each gene has been assigned a color. The arcs originate from different genes and terminate at different types of non-B DNA structures. The area of each colored ribbon depicts the frequency of that non-B DNA structure in the particular gene.

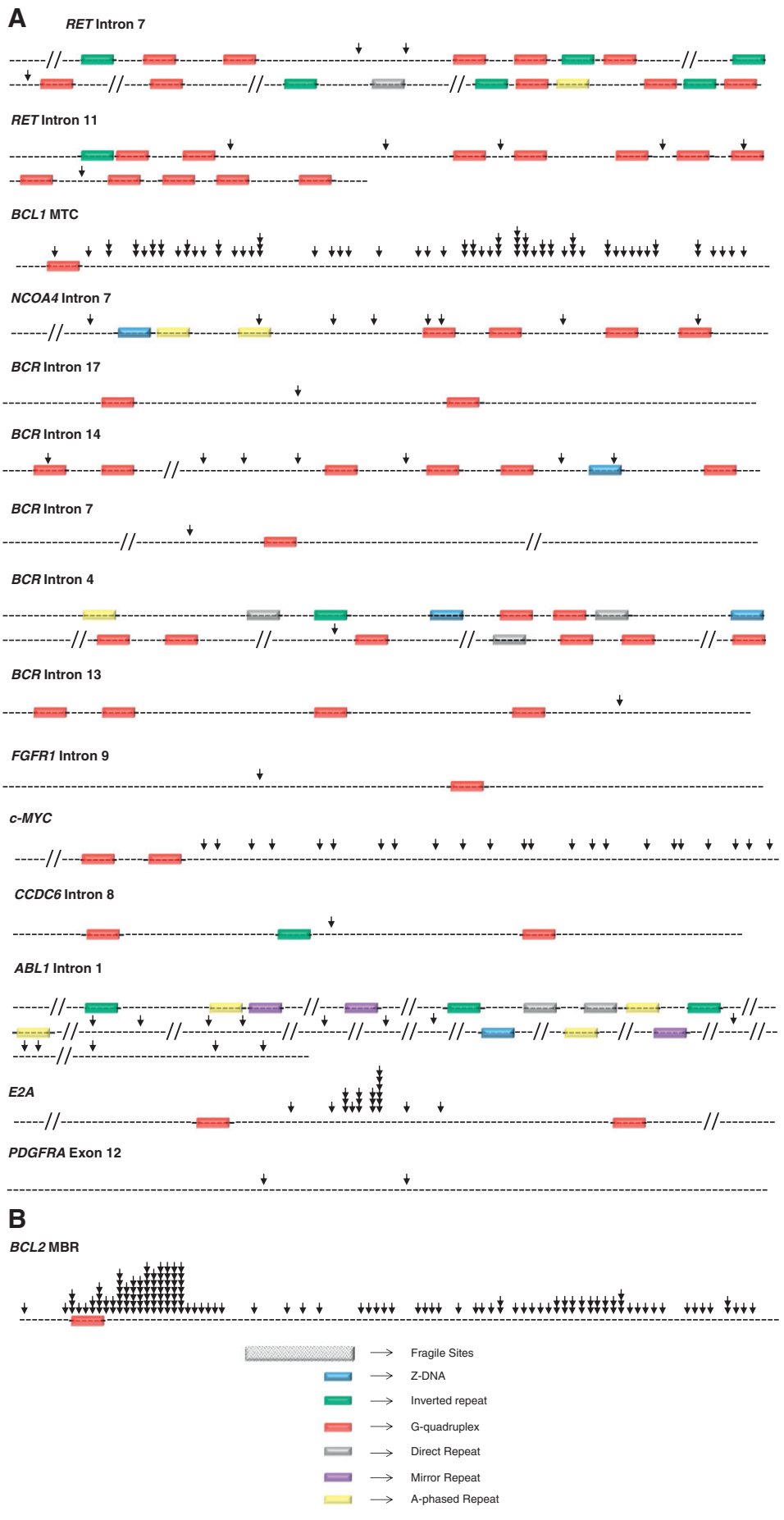
cluster region in at least 70% of the genes analyzed such as *E2A (TCF3)*, *BCR* intron 14, *HOX11* etc., which are involved in translocations in lymphoid cancers (Figs. 5A, B). Interestingly, such a correlation could not be drawn in case of the non-lymphoid cancers (Figs. 6A, B). Except *TMPRSS2* gene, which is involved in gene fusions in prostate cancers, none of the other genes like *ERG*, *FLI1*, *FUS*, and *EWSR1* showed the presence of G-quadruplex forming motifs within or at significant distances from the fragile regions (Fig. 6A). Overall these results suggest that among all the non-B DNA structures, G-quadruplexes are more frequently found near breakpoint regions in the genes engaged in chromosomal translocations.

3. Discussion

Chromosomal translocations are genetic hallmarks of cancers especially those of the haematopoeitic origin. The exact cause of why DNA breaks specifically in some regions leading to translocations is poorly understood. The incidence of such genomic rearrangements specifically in lymphoma and leukemia is generally shown to be due to the presence of the RAG proteins or the V(D)J recombinase which is expressed in the lymphoid cells in a stage-dependent manner [22,28–30]. The RAG complex is a nuclease with known sequence and structure specificity [21,31]. Studies have shown the involvement of RAGs in generating DNA double-strand breaks (DSBs) in both sequence and structure-specific dependent manner, in the context of chromosomal translocations [20,22,30,32–34]. Recent studies have implicated AID (*AICDA*) to be another potential mediator for generating translocations. AID (*AICDA*) was found to be responsible for creating DSBs during the *c-MYC (MYC)-IgH* translocations in Burkitt's lymphoma [35–38].

Since, RAG and AID (*AICDA*) proteins recognize altered DNA structures or single-strandedness in the DNA and lead to breaks, we wanted to test whether presence of non-B DNA structures influence genomic instability with respect to translocations. Using several bioinformatics

Fig. 3. Graphical representation of 16 hotspot genomic regions prone to breaks during chromosomal translocations in lymphoid cancers. A. Mapping of reported patient breakpoints and their proximity to different non-B DNA structures are shown. The arrows represent the position of the breakpoint junctions. The minimum and maximum distances between a non-B DNA structure and nearest breakpoint are 10 and 500 bp, respectively. B. Mapping of reported patient breakpoints for *BCL2* major breakpoint region (mbr) and its proximity to a G-quadruplex structure is shown. *BCL2* mbr is already shown in literature to form a G-quadruplex structure, hence validating our bioinformatics methods. Each colored box depicts a particular type of non-B DNA structure as indicated.


A
RET Intron 7
RET Intron 11
BCL1 MTC
NCOA4 Intron 7
BCR Intron 17
BCR Intron 14
BCR Intron 7
BCR Intron 4
BCR Intron 13
FGFR1 Intron 9
c-MYC
CCDC6 Intron 8
ABL1 Intron 1
E2A
PDGFRA Exon 12
B
BCL2 MBR
Fragile Sites
Z-DNA
Inverted repeat
G-quadruplex
Direct Repeat
Mirror Repeat
A-phased Repeat

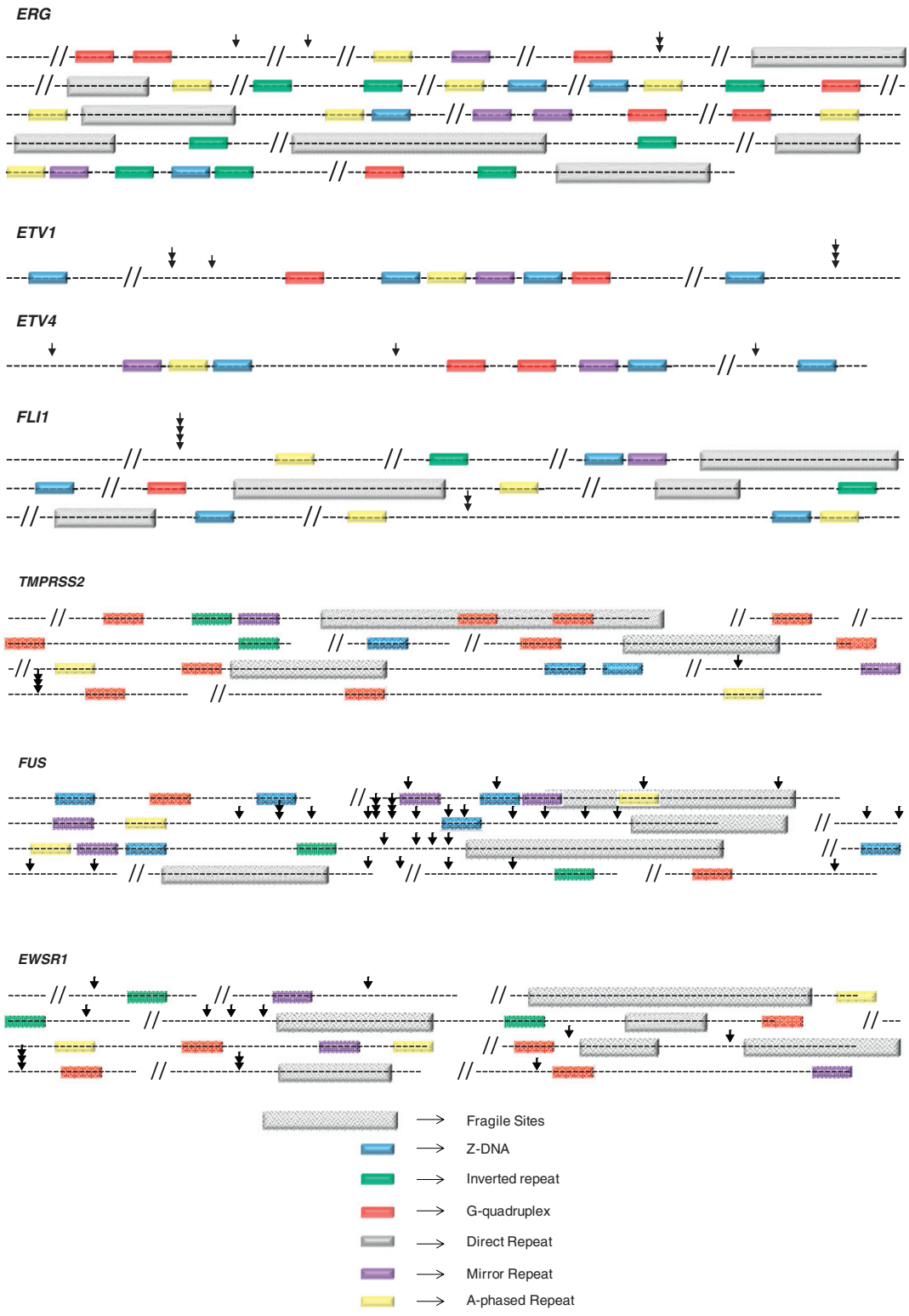


Fig. 4. Graphical representation of 7 genomic hotspot regions which are involved in chromosomal translocations in non-lymphoid cancers. The fragile regions and breakpoint positions reported in patients and their vicinity to various non-B DNA structures are depicted. Each colored box depicts a non-B DNA structure as indicated. The arrows represent the positions of breakpoint junctions.

tools we were able to identify the altered DNA structure formation in both *BCL2* mbr and *c-MYC (MYC)* genes, as reported previously [20,22]. Further, upon analyzing many other genes which are associated with translocations in both lymphoid and non-lymphoid cancers, we found several putative non-B DNA structure forming motifs in these breakpoint regions. We also observed a significantly high frequency of occurrence of such structures near many of the mapped patient breakpoint regions. More interestingly, the current study showed an interesting correlation between the presence of potential G-quadruplex forming motifs and frequency of the patient breakpoints near them. It is important to note that few recent studies have shown evidence for the importance of such G-quadruplex DNA in causing chromosomal

A

Sl. No.	Name of gene	Chr. No.	Name of Translocation	Length of breakpoint regions(bp)	No. of G-quadruplexes		Nearest G-quadruplex from 3G Sequence (bp)		Nearest G-quadruplex from 2G Sequence (bp)		Ensembl Sequence ID
					3Gs	2Gs	Upstream	Downstream	Upstream	Downstream	
1.	*BCL2* mbr	18	t(14,18)	150	1	6	~24300	~12000	~700	~700	60793432 - 60793581
2.	c-*MYC*	8	t(8,14)	1680	6	52	~1100	~350	0	0	128748528 - 128749549
3.	*BCL1* MTC	11	t(11,14)	170	0	3	~1600	~3800	0	~400	69346747 - 69346916
4.	E2A	19	t(1,19)	243	2	9	~2153	~500	~30	~30	1617706 - 1617948
5.	BCR Intron 14	22	t(9,22)	1281	2	19	~146	~510	~146	0	23632416 - 23633696
6.	*NCOA4* Intron 7	10	inv(10)(q11.2;q11.2)	976	0	12	~3024	~25000	~594	~409	51583492 - 51584467
7.	*HOX11*	10	t(10,14)	486	5	23	~4100	~2000	~200	~100	102890733 - 102891218

B

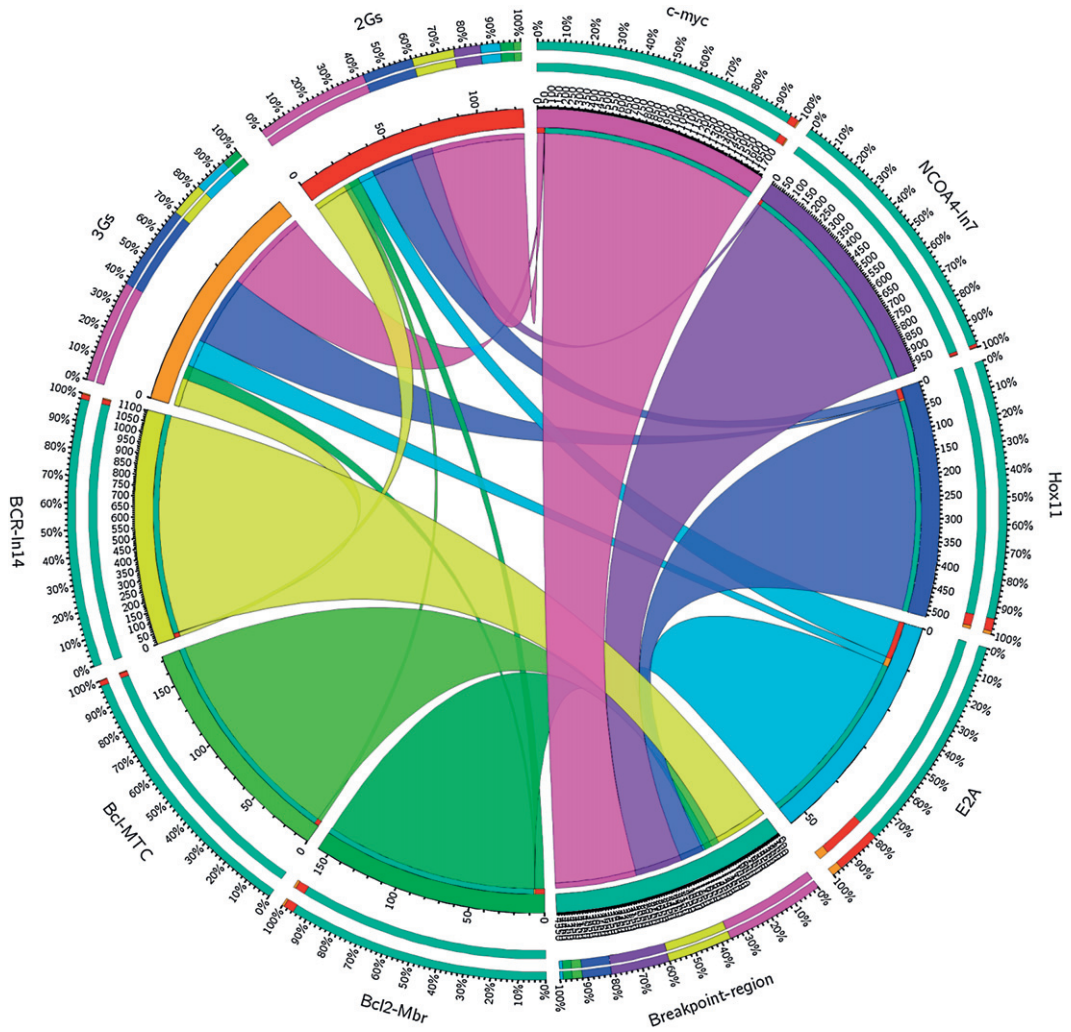


Fig. 5. Proximity of G-quadruplexes to the translocation breakpoint regions in lymphoid cancers. A. The table represents few selected genes, their major breakpoint regions and distance of the G-quadruplex structures in the vicinity of these fragile regions as seen in lymphoid cancers. The number of G-quadruplex forming motifs (3G) within the breakpoint regions is indicated (For *BCL2* mbr-1, *c-MYC (MYC)*-6, *E2A (TCF3)*, *BCR* intron 14–2 and *HOX11*-5). However, *BCL1 (CCND1)* MTC and *NCOA4* intron 7 do not show the presence of any 3G plate G-quadruplex structures. Distance of the next nearest G-quadruplex from the existing 3G or 2G quadruplex structures is also shown. B. Circos plot depicting the occurrence of G-quadruplex in genes involved in lymphoid cancers. Each gene is color coded and the area of the ribbons depicts the frequency of G-quadruplexes within these genes.

translocations. The *BCL2* mbr was shown to form a G-quadruplex DNA structure, susceptible to breaks by the RAG complex [20,22]. Similarly, the *c-MYC (MYC)* breakpoint region has also been suggested to form G-loop structures, which are responsible for their susceptibility to breaks [23]. Palindromic AT Rich Repeat (PATRR) associated genomic instability has been implicated in causing the recurrent t(11;22)(q23;q11) translocation [24,25,39]. In addition, cruciform DNA structure has been found to form at 11q23 in vitro and suggested to be responsible for translocations in this region [24,25,39]. Although non-lymphoid genes also had the propensity to form many non-B DNA structures near the fragile regions, we found limited evidence for the role of G-quadruplexes in causing fragility in these regions, unlike those of the lymphoid cancers. This suggests that in the non-lymphoid cancers some other mechanism could be playing a role in causing DNA breaks.

A

Sl. No	Name of gene	Chr No.	Name of Translocation	Length of breakpoint regions (bp)	No. of G-quadruplexes		Nearest G-quadruplex from 3G Sequence (bp)		Nearest G-quadruplex from 2G Sequence (bp)		Ensembl Sequence ID
					3Gs	2Gs	Upstream	Downstream	Upstream	Downstream	
1.	*ERG* Cluster 1	21	del(21)(q22.2-3)	540	0	1	~2090	~6204	~430	~201	39762869-39763408
2.	*ERG* Cluster 5	21	t(21,22)	1140	0	8	~6442	~4399	~494	~17	39947549-39948688
3.	*FLI1* Cluster 1	11	t(11,22)	1241	0	18	~7360	~2096	~3	~330	128642835–128644075
4.	*FLI1* Cluster 2	11	t(11,22)	1533	0	14	~1212	~9822	~130	~409	128651887-128653419
5.	*TMPRSS2* Cluster 1	21	inv(*21*)(*q22.2*;q22.*3*)	584	5	26	~111	~223	~0	~8	42879465-42880048
6.	*FUS* Cluster 1	16	t(12;16)(q13;p11)	1140	0	41	~653	~5499	~88	~432	31195585-31196724
7.	*EWSR1* Cluster 1	22	t(11,22)	1020	0	9	~5673	~1384	~483	~529	29683078-*29684097*

B

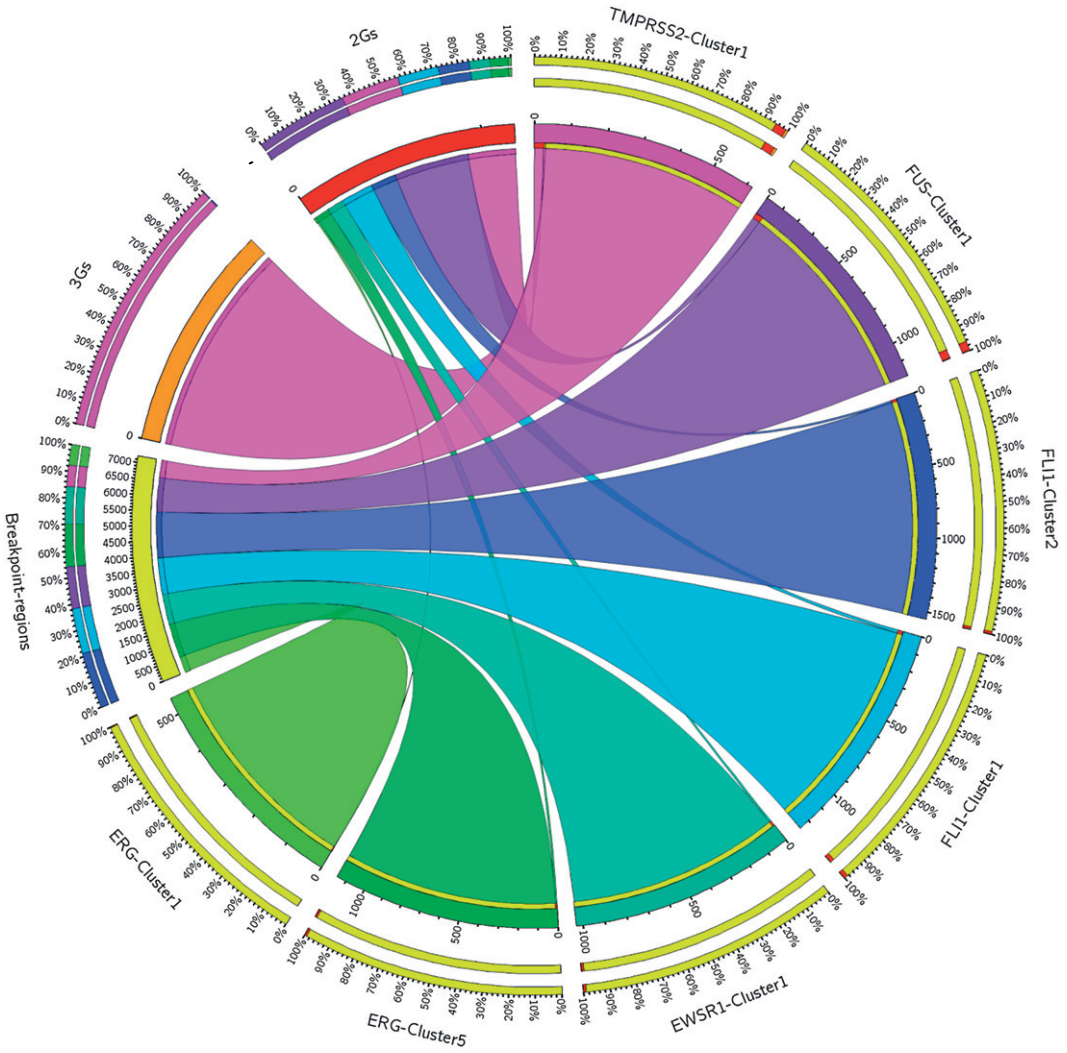


Fig. 6. Proximity of G-quadruplexes to the translocation breakpoint regions in non-lymphoid cancers. A. The table represents few selected genes, their major breakpoint regions and distance of the G-quadruplex structures in the vicinity of these fragile regions as seen in non-lymphoid cancers. Distance of nearest G-quadruplex from the hotspot region is mapped for both 2G and 3G plate quadruplex structures. B. Circos plot depicting the occurrence of G-quadruplex in genes involved in non-lymphoid cancers. Each gene is color coded and the area of the ribbons depicts the frequency of G-quadruplexes within these genes.

Alternatively, interaction between the other non-B DNA structures, flanking the fragile regions, could also impart instability to these regions.

Overall, this study demonstrates a direct correlation between the presence of non-B DNA structures, especially G-quadruplexes, and the occurrence of fragility of several genes involved during chromosomal translocations in cancers. These results will help in further deciphering the mechanism of generation of chromosomal translocations in many cancers which have not been studied so far. In addition, this study also describes few useful bioinformatic tools and approaches, which can be used to study the presence of altered DNA structures in the genome and its impact on genomic instability.

4. Materials and methods

4.1. Translocation breakpoint analysis

To analyze translocation breakpoint junctions, we retrieved data sets from specialized databases like TICdb and COSMIC [40–42]. TICdb has a widespread collection of translocation breakpoints which are finely mapped and has around 1445 different breakpoints from different reciprocal translocations. Entrez Gene was used to search genes of interest along with information of all the partner genes involved during the translocation. These are indicated by HGNC name of both partner genes with either their GenBank ID or Pubmed ID for the junction sequence. Further, for each partner gene, the fragment containing the breakpoint with an Ensembl transcript ID along with intron/exon number and the position of the breakpoint was obtained. The other database used in the study was COSMIC, which has curated information of somatic mutations in human cancers. This database has integrated Biomart which allows automated data mining and integration with other biological databases making higher end analysis possible. The gene name was entered in the search field and in the section "Rearrangement mutation summary" the name of the partner gene was selected to obtain complete information regarding the partner chromosome, number of patient samples analyzed and the total mutation junctions reported. cDNA sequence was retrieved and used for further mapping of breakpoints. With the above data, respective gene exon/intron sequence retrieval was carried out respectively using ENSEMBL and NCBI.

4.2. Non-B DNA structure analysis

To analyze formation of different types of non-B DNA structures in the retrieved gene sequence, a specialized database, non-B DB was used [26]. Non-B DB helps in integrating annotations and enables analysis of non-B DNA structure forming sequence motifs. Non-B DNA motif search tool (nBMST) was used wherein the non-B DNA motifs of interest were selected, their sequence added and the results obtained were used for further analysis. As non-B DB could not detect G-quadruplex with 2G repeats, another database named Quadbase [43] that is specially designed to analyze quadruplex motifs, was used. Pattern finder tool was used in our analysis by feeding the sequence data and adjusting the parameter to 2G or 3G.

In order to validate the results obtained from online database, Quadparser programme was used [44], which is a powerful computer algorithm to easily identify PQS from a large genomic data in FASTA format. In the command line interface (terminal on Mac OS X, command line on Windows), 'Quadparser' program was executed and the following parameters were used.

C:\xxxxxxxx\quadparser.exe <filename> <bases> <# bases in repeat> <repeats in sequence> <min gap size> <max gap size> <output file>

4.3. Mapping of translocation breakpoint and Non-B DNA structure

The breakpoint position and fragile region information obtained from TICdb and COSMIC databases were mapped on to the manually retrieved sequence. Different non-B DNA structures detected using non-B DB and Quadbase were also mapped onto the sequences retrieved.

4.4. Distance mapping of nearest G-quadruplex around breakpoint regions

The G-quadruplex forming sequence was mapped both downstream and upstream of the breakpoint cluster and the distance between them was calculated. G-quadruplex structures can be formed by repeats of 2G or more, and it is also known that for higher stability of the structure, a repeat of more than 3G is preferred. Therefore, we performed our analysis for both 2G and 3G repeats, separately. The results obtained by the analysis of the breakpoint cluster of selected genes were then tabulated. The data was further analyzed using Circos [45].

4.5. Circos analysis

Circos is a program used for visualizing data and information in a circular layout which makes it ideal for exploring relationships between positions or objects [45]. It assists the identification and analysis of relationships and differences arising from comparisons of genomes. This tool is used to show the variation in genome structure and any additional positional relationships between genomic intervals. The display of relationships between pairs of positions is facilitated by a circular ideogram layout by circos, with the use of ribbons, which encode the position, size, and orientation of related genomic elements [45]. It was primarily designed for displaying genomic data, mainly cancer genomics and comparative genomics and molecular biology. With its special features it addresses classic challenges in representing such data, which tend to be very sparse and encompass a large number of length scales.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.ygeno.2012.05.008.

Acknowledgments

We would like to thank Divyaanka, Nishana and other members of SCR laboratory for help, discussions and critical reading of the manuscript. This work was supported by grants from CSIR, India 37(1400/10/EMR-II) for SCR. MN is supported by postdoctoral fellowship of IISc, Bangalore, India.

References

[1] J.D. Rowley, Chromosome translocations: dangerous liaisons revisited, Nat. Rev. Cancer 1 (2001) 245–250.
[2] M. Nambiar, V. Kari, S.C. Raghavan, Chromosomal translocations in cancer, Biochim. Biophys. Acta 1786 (2008) 139–152.
[3] A. Nussenzweig, M.C. Nussenzweig, Origin of chromosomal translocations in lymphoid cancer, Cell 141 (2010) 27–38.
[4] T.H. Rabbitts, Chromosomal translocations in human cancer, Nature 372 (1994) 143–149.
[5] S.J. Korsmeyer, Chromosomal translocations in lymphoid malignancies reveal novel proto-oncogenes, Annu. Rev. Immunol. 10 (1992) 785–807.
[6] C. Kumar-Sinha, S.A. Tomlins, A.M. Chinnaiyan, Recurrent gene fusions in prostate cancer, Nat. Rev. Cancer 8 (2008) 497–511.
[7] S.A. Tomlins, D.R. Rhodes, S. Perner, S.M. Dhanasekaran, R. Mehra, X.W. Sun, S. Varambally, X. Cao, J. Tchinda, R. Kuefer, C. Lee, J.E. Montie, R.B. Shah, K.J. Pienta, M.A. Rubin, A.M. Chinnaiyan, Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer, Science 310 (2005) 644–648.
[8] S.A. Tomlins, B. Laxman, S.M. Dhanasekaran, B.E. Helgeson, X. Cao, D.S. Morris, A. Menon, X. Jing, Q. Cao, B. Han, J. Yu, L. Wang, J.E. Montie, M.A. Rubin, K.J. Pienta, D. Roulston, R.B. Shah, S. Varambally, R. Mehra, A.M. Chinnaiyan, Distinct classes of chromosomal rearrangements create oncogenic ETS gene fusions in prostate cancer, Nature 448 (2007) 595–599.
[9] M.L. Cleary, J. Sklar, Nucleotide sequence of a t(14;18) chromosomal breakpoint in follicular lymphoma and demonstration of a breakpoint-cluster region near a transcriptionally active locus on chromosome 18, Proc. Natl. Acad. Sci. U. S. A. 82 (1985) 7439–7443.
[10] A. Bakhshi, J.P. Jensen, P. Goldman, J.J. Wright, O.W. McBride, A.L. Epstein, S.J. Korsmeyer, Cloning the chromosomal breakpoint of t(14;18) human lymphomas: clustering around JH on chromosome 14 and near a transcriptional unit on 18, Cell 41 (1985) 899–906.
[11] A. Bakhshi, J.J. Wright, W. Graninger, M. Seto, J. Owens, J. Cossman, J.P. Jensen, P. Goldman, S.J. Korsmeyer, Mechanism of the t(14;18) chromosomal translocation: structural analysis of both derivative 14 and 18 reciprocal partners, Proc. Natl. Acad. Sci. U. S. A. 84 (1987) 2396–2400.
[12] J.D. Rowley, A new consistent chromosomal abnormality in chronic Myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining, Nature 243 (1973) 290–293.
[13] P.C. Nowell, D.A. Hungerford, A minute chromosome in human chronic granulocytic leukemia, Science 132 (1960) 1497.
[14] R.R. Sinden (Ed.), DNA Structure and Function, Academic Press, San Diego, California, 1994.
[15] R.R. Sinden, Biological implications of the DNA structures associated with disease-causing triplet repeats, Am. J. Hum. Genet. 64 (1999) 346–353.

[16] S.C. Raghavan, M.R. Lieber, DNA structures at chromosomal translocation sites, Bioessays 28 (2006) 480–494.
[17] A. Bacolla, R.D. Wells, Non-B DNA conformations, genomic rearrangements, and human disease, J. Biol. Chem. 279 (2004) 47411–47414.
[18] S. Neidle, S. Balasubramanian, Quadruplex Nucleic Acids, RSC Biomolecular Sciences, 2006.
[19] A.K. Todd, M. Johnston, S. Neidle, Highly prevalent putative quadruplex sequence motifs in human DNA, Nucleic Acids Res. 33 (2005) 2901–2907.
[20] M. Nambiar, G. Goldsmith, B.T. Moorthy, M.R. Lieber, V.M. Joshi, B. Choudhary, R.V. Hosur, S.C. Raghavan, Formation of a G-quadruplex at the BCL2 major breakpoint region of the t(14;18) translocation in follicular lymphoma, Nucleic Acids Res. 39 (2011) 936–948.
[21] M. Nambiar, S.C. Raghavan, How does DNA break during chromosomal translocations? Nucleic Acids Res. 39 (2011) 5813–5825.
[22] S.C. Raghavan, P.C. Swanson, X. Wu, C.L. Hsieh, M.R. Lieber, A non-B-DNA structure at the Bcl-2 major breakpoint region is cleaved by the RAG complex, Nature 428 (2004) 88–93.
[23] M.L. Duquette, P. Pham, M.F. Goodman, N. Maizels, AID binds to transcription-induced structures in c-MYC that map to regions associated with translocation and hypermutation, Oncogene. 24 (2005) 5791–5798.
[24] H. Kurahashi, H. Inagaki, T. Ohye, T. Kogo, T. Kato, B.S. Emanuel, Palindrome-mediated chromosomal translocations in humans, DNA Repair (Amst) 5 (2006) 1136–1145.
[25] H. Kurahashi, H. Inagaki, K. Yamada, T. Ohye, M. Taniguchi, B.S. Emanuel, T. Toda, Cruciform DNA structure underlies the etiology for palindrome-mediated human chromosomal translocations, J. Biol. Chem. 279 (2004) 35377–35383.
[26] R.Z. Cer, K.H. Bruce, U.S. Mudunuri, M. Yi, N. Volfovsky, B.T. Luke, A. Bacolla, J.R. Collins, R.M. Stephens, Non-B DB: a database of predicted non-B DNA-forming motifs in mammalian genomes, Nucleic Acids Res. 39 (2011) D383–D391.
[27] A. Bacolla, A. Jaworski, J.E. Larson, J.P. Jakupciak, N. Chuzhanova, S.S. Abeysinghe, C.D. O'Connell, D.N. Cooper, R.D. Wells, Breakpoints of gross deletions coincide with non-B DNA conformations, Proc. Natl. Acad. Sci. U. S. A. 101 (2004) 14162–14167.
[28] M. Nambiar, S.C. Raghavan, Mechanism of fragility at BCL2 gene minor breakpoint cluster region during t(14;18) chromosomal translocation, J. Biol. Chem. 287 (2012) 8688–8701.
[29] S.C. Raghavan, I.R. Kirsch, M.R. Lieber, Analysis of the V(D)J recombination efficiency at lymphoid chromosomal translocation breakpoints, J. Biol. Chem. 276 (2001) 29126–29133.
[30] R. Marculescu, T. Le, P. Simon, U. Jaeger, B. Nadel, V(D)J-mediated translocations in lymphoid neoplasms: a functional assessment of genomic instability by cryptic sites, J. Exp. Med. 195 (2002) 85–98.
[31] S.D. Fugmann, A.I. Lee, P.E. Shockett, I.J. Villey, D.G. Schatz, The RAG proteins and V(D)J recombination: complexes, ends, and transposition, Annu. Rev. Immunol. 18 (2000) 495–527.
[32] M. Zhang, P.C. Swanson, V(D)J recombinase binding and cleavage of cryptic recombination signal sequences identified from lymphoid malignancies, J. Biol. Chem. 283 (2008) 6717–6727.
[33] S.M. Lewis, E. Agard, S. Suh, L. Czyzyk, Cryptic signals and the fidelity of V(D)J joining, Mol. Cell. Biol. 17 (1997) 3125–3136.
[34] S.C. Raghavan, P.C. Swanson, Y. Ma, M.R. Lieber, Double-strand break formation by the RAG complex at the bcl-2 major breakpoint region and at other non-B DNA structures in vitro, Mol. Cell. Biol. 25 (2005) 5904–5919.
[35] D.F. Robbiani, A. Bothmer, E. Callen, B. Reina-San-Martin, Y. Dorsett, S. Difilippantonio, D.J. Bolland, H.T. Chen, A.E. Corcoran, A. Nussenzweig, M.C. Nussenzweig, AID is required for the chromosomal breaks in c-myc that lead to c-myc/IgH translocations, Cell 135 (2008) 1028–1038.
[36] D.F. Robbiani, S. Bunting, N. Feldhahn, A. Bothmer, J. Camps, S. Deroubaix, K.M. McBride, I.A. Klein, G. Stone, T.R. Eisenreich, T. Ried, A. Nussenzweig, M.C. Nussenzweig, AID produces DNA double-strand breaks in non-Ig genes and mature B cell lymphomas with reciprocal chromosome translocations, Mol. Cell 36 (2009) 631–641.
[37] Y. Dorsett, M.J. Robbiani, Bernardo Reina-San-Martin, Thomas R. Eisenreich, Michel C. Nussenzweig, A role for AID in chromosome translocations between c-myc and the IgH variable region, J. Exp. Med. 204 (2007) 2225–2232.
[38] A.R. Ramiro, M. Jankovic, T. Eisenreich, S. Difilippantonio, S. Chen-Kiang, M. Muramatsu, T. Honjo, A. Nussenzweig, M.C. Nussenzweig, AID is required for c-myc/IgH chromosome translocations in vivo, Cell 118 (2004) 431–438.
[39] H. Kurahashi, T.H. Shaikh, P. Hu, B.A. Roe, B.S. Emanuel, M.L. Budarf, Regions of genomic instability on 22q11 and 11q23 as the etiology for the recurrent constitutional t(11;22), Hum. Mol. Genet. (2000) 9.
[40] F.J. Novo, I.O. de Mendibil, J.L. Vizmanos, TICdb: a collection of gene-mapped translocation breakpoints in cancer, BMC Genomics 8 (2007) 33.
[41] S.A. Forbes, G. Tang, N. Bindal, S. Bamford, E. Dawson, C. Cole, C.Y. Kok, M. Jia, R. Ewing, A. Menzies, J.W. Teague, M.R. Stratton, P.A. Futreal, COSMIC (the Catalogue of Somatic Mutations in Cancer): a resource to investigate acquired mutations in human cancer, Nucleic Acids Res. 38 (2010) D652–D657.
[42] S.A. Forbes, N. Bindal, S. Bamford, C. Cole, C.Y. Kok, D. Beare, M. Jia, R. Shepherd, K. Leung, A. Menzies, J.W. Teague, P.J. Campbell, M.R. Stratton, P.A. Futreal, COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer, Nucleic Acids Res. 39 (2011) D945–D950.
[43] V.K. Yadav, J.K. Abraham, P. Mani, R. Kulshrestha, S. Chowdhury, QuadBase: genome-wide database of G4 DNA—occurrence and conservation in human, chimpanzee, mouse and rat promoters and 146 microbes, Nucleic Acids Res. 36 (2008) D381–D385.
[44] J.L. Huppert, S. Balasubramanian, Prevalence of quadruplexes in the human genome, Nucleic Acids Res. 33 (2005) 2908–2916.
[45] M. Krzywinski, J. Schein, I. Birol, J. Connors, R. Gascoyne, D. Horsman, S.J. Jones, M.A. Marra, Circos: an information aesthetic for comparative genomics, Genome Res. 19 (2009) 1639–1645.